1990

Allozyme Diversity and Introgression in the Galapagos Islands Endemic Gossypium darwinii and its Relationship to Continental G. barbadense

Jonathan F. Wendel
Iowa State University, jfw@iastate.edu

Richard G. Percy
United States Department of Agriculture

Biochemical Systematics and Ecology, Vol. 18, No. 7/8, pp. 517–528, 1990.
Printed in Great Britain.

0305-1978/90 $3.00+0.00

Allozyme Diversity and Introgression in the Galapagos Islands Endemic *Gossypium darwinii* and its Relationship to Continental *G. barbadense*

JONATHAN F. WENDEL* and RICHARD G. PERCY†

*Department of Botany, Bessey Hall, Iowa State University, Ames, IA 50011, U.S.A.;
†USDA-ARS, in cooperation with the Arizona Agricultural Experimental Station, Maricopa Agricultural Center, 37860 W. Smith-Enke Rd, Maricopa, AZ 85339, U.S.A.

Key Word Index—*Gossypium darwinii*; *G. barbadense*; Malvaceae; cotton; introgression; Galapagos Islands; genetic variation; island evolution; allozymes; molecular divergence.

Abstract—*Gossypium darwinii* Watt is a tetraploid cotton endemic to the Galapagos Islands. Opinion has been divided as to whether or not it deserves recognition at the specific rank, with some considering it a variety of its presumed progenitor, the widely distributed South American species *G. barbadense* L. A previous hypothesis states that much of the perceived intergradation between the two taxa arose as a consequence of introgression from *G. barbadense* following its introduction to the archipelago during the past several hundred years. We performed allozyme analysis on 58 accessions of *G. darwinii* from six islands, using 17 enzymes collectively encoded by 59 loci. Levels of variation were high for an island endemic, with a mean number of alleles per locus of 1.34 and an average panmictic heterozygosity of 0.062. Principal component analysis revealed clustering of accessions according to their island of origin, and a spatial pattern of island-clusters that approximates geographical relationships among islands. Genetic relationships of *G. darwinii* with *G. barbadense* and *G. hirsutum* L. were studied using previously generated allozyme data. Significant introgression of *G. hirsutum* alleles was detected; however morphological considerations support the hypothesis that much of *G. darwinii*'s diversity stems from interspecific gene flow from *G. barbadense*. Evidence is presented suggesting that the occurrence of *G. hirsutum* alleles in *G. darwinii* derives not from direct hybridization, but from a mediated transfer through introduced, *G. hirsutum*-introgressed *G. barbadense*. *Gossypium darwinii* and *G. barbadense* are nearly fixed for different alleles at four loci and each contains a large number of unique alleles. Notwithstanding the high interspecific Nei's genetic identity (0.949), the allozyme data support geographical and morphological evidence in suggesting that a specific rank for *G. darwinii* is warranted.

Introduction

Approximately 500 species of vascular plants are indigenous to the Galapagos Islands, including about 200 endemics [1–3]. Two of these endemics are members of the cotton genus (*Gossypium* L.), *G. klotzschianum* Anderss. (reported from five of the islands) and *G. darwinii* Watt , (from 13 islands). Phytogeographic evidence suggests that most of the Galapagos flora (approximately 90%) arose from western South American ancestors, some 1000 km or more distant [1–3]. An exception to this generalization is *G. klotzschianum*, one of 13 diploid ($2n=26$) *Gossypium* species native to the Western Hemisphere. *Gossypium klotzschianum* is unique among the Galapagos endemics in that it apparently originated from a northern Mexican progenitor [4]. In contrast, the closest relative of *G. darwinii*, and its presumed progenitor, is *G. barbadense* L., one of two commercially important (with *G. hirsutum* L.) tetraploid ($2n=4x=52$) species of cotton. *Gossypium barbadense* has an indigenous cultivated range throughout the New World tropics, but its pre-agricultural distribution appears to be northwestern S. American [5, 6].

Opinion has been divided as to whether *G. darwinii* deserves recognition at the specific rank. Watt, in honoring the Galapagos' most famous visitor with the commemorative epithet, emphasized several morphological discontinuities between the two taxa and their geographic isolation [7]. He stated that "At the time at which this species was collected by Darwin [in 1835 on San Salvador, including the type specimen] . . . cotton could hardly have been cultivated anywhere on the Galapagos Islands . . .", and concluded that "it is a perfectly good

(*Received* 25 *June* 1990)

species." Later adherents of this view include Harland [8], based on comparative genetic analyses, and Fryxell [9], in the most recent monograph of the genus. The alternative opinion was promulgated by Hutchinson *et al.* [5], who reduced *G. darwinii* to varietal rank [*G. barbadense* var. *darwinii* (Watt) J. B. Hutch.]. Kearney [10] and Valíček [11] concurred, the latter author stating that *G. darwinii* is "... decidedly not an independent (separate) species."

Hutchinson *et al.* [5] offered several reasons for reducing *G. darwinii* to varietal status, including its highly variable morphology, the vigorous and fertile F_2 generation obtained from interspecific *G. barbadense* × *G. darwinii* hybrids, and the absence of any "clear line of demarcation, either morphological or geographical, between the various forms of *darwinii* and *barbadense*." However, they further suggested that the perceived intergradation may be a consequence of the introduction of cultivated *G. barbadense* from the mainland during the last two centuries and its subsequent hybridization and introgression with native *G. darwinii*. In our observations of numerous accessions of both species over several years, we have noted a relatively high frequency of "improved" characteristics (e.g. longer fiber and larger capsules) in *G. darwinii*, lending credence to the suggestion that species boundaries have been obscured by historical, interspecific introgression.

The objective of the present study, was to assess, by means of allozyme analysis, the genetic relationships between *G. darwinii* and *G. barbadense*. Specifically, we wished to: (i) quantify the amount of genetic variability within *G. darwinii* and its patterns of inter-island differentiation; (ii) elucidate the magnitude of the "genetic challenge" posed by interspecific introgression from *G. barbadense*, and assess the significance of this gene flow to the level and pattern of diversity presently observed in *G. darwinii*; and (iii) test the hypothesis that *G. barbadense* and *G. darwinii* represent a progenitor-derivative species pair.

Results

Geographic patterns and genetic variability

Seventeen enzymes were screened for electrophoretic variation in 58 accessions originating from throughout the species' range (Table 1, Fig. 1). Evidence from previous studies indicated that these 17 enzymes are encoded by a minimum of 59 genetic loci [6]. Allelic variants at 14 of these loci distinguish *G. barbadense* from *G. hirsutum* [6; Wendel, unpublished data], providing the capability of documenting interspecific introgression (see *Discussion*). An initial inspection of the allelic presence data for the 58 *G. darwinii* accessions led to the recognition of seven accessions (AS892, CB3087, CB3097, CB3098, WB1207, WB1249, PW56) that contained one or more *G. hirsutum* alleles. These seven accessions were dropped from the initial data set and are not included in the results presented in Tables 2–5.

For the remaining 51 accessions, 43 of 59 loci were fixed for the same allele. At least one locus was variable for nine enzyme systems, resulting in a total of 16 polymorphic loci (P=27.1% for the species) and 36 allelic variants (Table 2). Including the 43 monomorphic loci, the mean number of alleles per locus in *G. darwinii* is 1.34 (2.25 per polymorphic locus). The majority of polymorphic loci are only weakly polymorphic; frequencies of the most common alleles are equal to or greater

TABLE 1. ACCESSIONS* INCLUDED IN ALLOZYME STUDY OF *GOSSYPIUM DARWINII*, ARRANGED BY ISLAND OF ORIGIN

Island	Accessions studied
Eden	AS989, WB1200
Gardner	AS899, AS911, AS914, CB3097
Floreana	AS883, AS885, AS892, AS893, AS894, AS896, CB3098, CB3120, WB1207
Isabela	PW22, PW23, PW29, PW30, PW32, PW33, PW36, PW37, PW38, PW44, PW45, PW49, PW50, PW51, PW53, PW56, WB1203, WB1242, WB1249
San Cristóbal	AS926, AS953, AS955, AS956, AS958, AS968, AS970, CB3087, CB3973†, WB1215, WB1225, WB1227, WB1229, WB1239†, WB1245
Santa Cruz	PW2, PW5, PW6, PW7, PW9, PW10, PW11, PW12, WB1201

*Accessions preceded by the letters WB are from a collection assembled by S. G. Stephens at North Carolina State University, Raleigh, NC; those preceded by the letters CB were part of a collection formerly maintained at the "Cotton and Cordage Fibers Research Branch" of the USDA-ARS National Headquarters in Beltsville, MD (now stored in the National Germplasm Collection at Fort Collins, CO); accessions preceded by PW were collected by Percival and Wilson in 1985 [37], and those preceded by AS were collected by Schwendiman *et al.* in 1983 [38]; all accessions are presently maintained in a working collection at Maricopa, AZ.

†These two accessions are from the small island of Lobos adjacent to San Cristóbal.

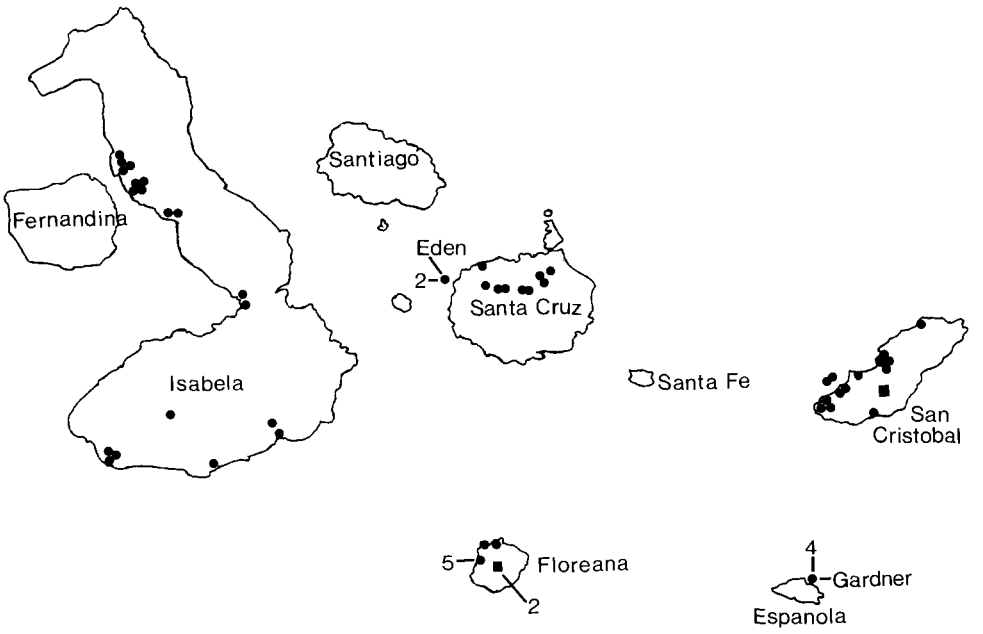

FIG. 1. MAP OF A PORTION OF THE GALAPAGOS ISLANDS SHOWING COLLECTION LOCALITIES FOR *GOSSYPIUM DARWINII* ACCESSIONS. Only island of origin is known for several accessions (closed squares); more precise locality information exists for most accessions (closed circles). Arabic numerals indicate the number of accessions that map to the same point at this scale of resolution.

than 0.90 for eight of the 16 polymorphic loci (Table 2). Consequently, most loci have relatively low estimates of panmictic heterozygosity. Averaged across polymorphic loci, the mean panmictic heterozygosity in *G. darwinii* is 0.230; including the 43 monomorphic loci, it is 0.062.

In agreement with preliminary observations and published data for other *Gossypium* species [4, 6, 12, 13] most populations of *G. darwinii* appear to be fixed for a single multi-locus genotype, although low levels of variation were sometimes observed. In the majority of cases where variation was observed within a population, it was limited to only one or two loci. Similarly, observed heterozygosity was nearly non-existent; when an accession was polymorphic, it usually consisted of a mixture of alternate homozygotes.

Most of the 36 alleles detected at the polymorphic loci appeared to have non-random distributions among the six islands sampled. Geographic patterns of allelic distribution varied widely, with seven alleles restricted in occurrence to individual islands and many others confined to various combinations of islands (Tables 2, 3). For example, *Pgm6-9* reaches a moderate to high frequency on Santa Cruz, Gardner and San Cristóbal but was not detected from the other three islands. Although many different geographic patterns of distribution were observed, there appeared to be a relationship between geographic proximity and allelic occurrence, i.e. nearby islands were more likely to share alleles than more distant islands. Chi-square tests of gene frequency homogeneity [14] indicated that gene frequencies for 12 of the 16 polymorphic loci are significantly heterogeneous among the six islands ($p<0.05$; Table 4). Gene diversity statistics closely parallel these results; the proportion of genetic variation resulting from differences among regions (G_{ST}) ranged from negligible for *Tpi7* (equivalent gene frequencies in all islands) to 0.65 for *Aco3*. Averaged over the 16 polymorphic loci, the proportion of total variation arising as a consequence of differentiation among islands was 0.43.

Principal component analysis (PCA) was performed on the covariance matrix of allele frequencies. Accessions were plotted according to their coordinates along the first two principal

TABLE 2. ALLELE FREQUENCIES* AT 23 ALLOZYME LOCI IN *GOSSYPIUM DARWINII* (BY ISLAND) AND *G. BARBADENSE*

Locus	Allele	Santa Cruz (9)	Gardner (4)	Isabela (19)	Floreana (9)	San Cristóbal (15)	Eden (2)	*G. darwinii* (51)	*G. barbadense* (111)
Adh2	*1*	1.00	1.00	0.65	0.50	1.00	1.00	0.82	0.00
	4	0.00	0.00	0.35	0.50	0.00	0.00	0.18	1.00
Mdh4	*4*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.32
	6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.68
Gdh4	*4*	0.06	1.00	0.00	0.67	0.29	0.00	0.23	0.42
	6	0.94	0.00	1.00	0.33	0.71	1.00	0.77	0.58
Idh1	*1*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97
Idh2	*4*	0.89	0.67	0.71	0.33	1.00	1.00	0.78	0.49
	5	0.11	0.33	0.29	0.67	0.00	0.00	0.22	0.51
Enp1	*4*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12
	5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.88
Enp2	*4*	1.00	1.00	0.38	1.00	1.00	0.25	0.77	0.99
	5	0.00	0.00	0.62	0.00	0.00	0.75	0.23	0.00
	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Tpi3	*4*	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Tpi6	*4*	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Tpi7	4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	5	0.89	0.33	1.00	1.00	0.93	1.00	0.92	0.99
	6	0.11	0.67	0.00	0.00	0.07	0.00	0.08	0.00
Arg2	*0.2*	0.09	0.00	0.00	0.00	0.00	0.00	0.02	0.00
	0.5	0.91	1.00	0.82	0.92	1.00	1.00	0.91	0.00
	3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	4	0.00	0.00	0.18	0.08	0.00	0.00	0.07	0.89
	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09
	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Aat2	*2*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.96
Pgd1	*1*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97
Aco1	*2*	0.33	0.00	0.00	0.00	0.00	0.00	0.06	0.00
	4	0.67	1.00	1.00	1.00	1.00	1.00	0.94	0.92
	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04
	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04
Aco3	*1*	0.11	0.00	0.00	0.00	0.00	0.00	0.02	0.01
	4	0.44	0.00	0.82	0.00	0.07	1.00	0.41	0.92
	7	0.44	1.00	0.18	1.00	0.93	0.00	0.57	0.02
	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05
Aco5	*1*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
	4	1.00	1.00	1.00	1.00	0.93	1.00	0.98	0.95
	8	0.00	0.00	0.00	0.00	0.07	0.00	0.02	0.00
	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
Aco6	*2*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	4	1.00	1.00	0.82	1.00	1.00	1.00	0.94	0.91
	9	0.00	0.00	0.18	0.00	0.00	0.00	0.06	0.08
Leu1	*2*	0.00	0.00	0.00	0.33	0.00	0.50	0.06	0.04
	3	1.00	1.00	0.65	0.33	0.93	0.50	0.77	0.00
	4	0.00	0.00	0.35	0.33	0.00	0.00	0.16	0.96
	9	0.00	0.00	0.00	0.00	0.07	0.00	0.02	0.01
Pgm1	*2*	0.00	0.00	0.06	0.00	0.00	0.00	0.02	0.00
	4	1.00	1.00	0.94	1.00	1.00	1.00	0.98	0.81
	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19
Pgm4	*4*	0.94	1.00	0.65	1.00	1.00	1.00	0.87	0.85
	9	0.06	0.00	0.35	0.00	0.00	0.00	0.13	0.15
Pgm5	*2*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
	4	0.94	1.00	1.00	1.00	1.00	1.00	0.99	0.97

TABLE 2—*CONTINUED*

		Santa Cruz (9)	Gardner (4)	Isabela (19)	Floreana (9)	San Cristóbal (15)	Eden (2)	*G. darwinii* (51)	*G. barbadense* (111)
	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
	9	0.06	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Pgm6	*4*	0.11	0.33	1.00	1.00	0.68	1.00	0.72	1.00
	9	0.89	0.67	0.00	0.00	0.32	0.00	0.28	0.00
Pgm7	*4*	1.00	1.00	0.94	1.00	0.71	1.00	0.90	0.70
	8	0.00	0.00	0.06	0.00	0.29	0.00	0.10	0.00
	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30

*Island frequencies are unweighted arithmetic means of population frequencies (number of populations in parentheses). Seven *G. darwinii* accessions were removed from the data set because they contained introgressed *G. hirsutum* germplasm. Data for *G. barbadense* represent non-introgressant accessions from Percy and Wendel [6].

TABLE 3. SUMMARY STATISTICS OF GENETIC VARIABILITY AT 59 LOCI IN *GOSSYPIUM DARWINII* AND *G. BARBADENSE**

	Santa Cruz (9)	Gardner (4)	Isabela (19)	Floreana (9)	San Cristóbal (15)	Eden (2)	*G. darwinii* (51)	*G. barbadense* (111)
No of unique alleles†	4	0	2	0	1	0	12	21
Polymorphic loci (%)	15.3	5.1	16.9	8.5	11.9	3.4	27.1	35.6
Mean no. of alleles per locus	1.17	1.04	1.17	1.08	1.13	1.04	1.34	1.51
Mean panmictic heterozygosity	0.033	0.023	0.057	0.037	0.030	0.015	0.062	0.064

*Includes the 16 loci that were polymorphic in *G. darwinii* (from Table 2) and 43 additional monomorphic loci. Number of populations per island is in parentheses. Seven *G. darwinii* accessions were removed from the data set because they contained introgressed *G. hirsutum* germplasm. *Gossypium barbadense* estimates were derived using non-introgressant accessions from Percy and Wendel [6].

†Defined as the number of alleles restricted to an island, or in the case of *G. barbadense* v *G. darwinii*, the number of alleles confined to one of the two species.

TABLE 4. GENETIC DIFFERENTIATION* OF *GOSSYPIUM DARWINII* POPULATIONS AMONG SIX GALAPAGOS ISLANDS FOR 16 ALLOZYME LOCI

Locus	H_S	H_T	G_{ST}	$X^2(df)$	*P*
Adh2	0.16	0.30	0.47	27.9 (5)	0.00
Gdh1	0.16	0.34	0.53	48.6 (5)	0.00
Idh2	0.28	0.34	0.19	26.1 (5)	0.00
Enp2	0.15	0.36	0.59	53.2 (5)	0.00
Tpi7	0.15	0.15	0.02	33.2 (5)	0.00
Arg2	0.10	0.13	0.30	10.3 (5)	0.06
Aco1	0.08	0.11	0.29	29.7 (5)	0.00
Aco3	0.18	0.51	0.65	66.2 (10)	0.00
Aco5	0.02	0.04	0.37	5.3 (5)	0.37
Aco6	0.05	0.11	0.55	12.7 (5)	0.02
Leu1	0.31	0.39	0.21	65.0 (15)	0.00
Pgm1	0.02	0.04	0.37	4.0 (5)	0.53
Pgm4	0.10	0.22	0.55	23.6 (5)	0.00
Pgm5	0.01	0.02	0.50	4.7 (5)	0.45
Pgm6	0.20	0.41	0.51	56.6 (5)	0.00
Pgm7	0.09	0.18	0.50	16.0 (5)	0.01
Mean	0.13	0.23	0.43		
Mean with 43 invariant loci	0.03	0.06			

*H_S, H_T and G_{ST} are gene diversty statistics of Nei [22]. The final two columns present chi-square values from tests of gene frequency homogeneity [14] among islands (*df*) and their probability values, under the null hypothesis of no genetic differentiation.

components (which accounted for 43.7% of the total variance). The resulting plot led to the recognition of reasonably discrete clusters of accessions that correspond to particular islands (Fig. 2). The first principal component (PCA1) separated most Isabela populations (large positive PCA1 scores) from those collected on San Cristóbal and Gardner (large negative PCA1 scores); populations from the remaining islands had intermediate PCA1 scores. PCA2 separates these latter populations into clusters representing Santa Cruz (large positive PCA2 scores) and Floreana (large negative PCA2 scores).

Overall affinity of populations from different islands was summarized by measures of genetic similarity (Table 5). Nei's genetic identity (*I*) ranged from a minimum of 0.939 for the Eden–Gardner pair to a maximum of 0.992 for the Eden–Isabela pair (mean inter-island $\bar{I}$=0.969). These results were expected from the PCA analysis (Fig. 2). Other estimates of *I* were also in accordance with expectations, particularly the close relationship (high estimates of *I*) among

TABLE 5. GENETIC RELATIONSHIP AMONG *GOSSYPIUM DARWINII* ACCESSIONS FROM DIFFERENT GALAPAGOS ISLANDS*

	Santa Cruz	Gardner	Isabela	Floreana	San Cristóbal	Eden
Santa Cruz	—	0.974	0.969	0.960	0.988	0.969
Gardner	0.026	—	0.943	0.974	0.982	0.939
Isabela	0.031	0.058	—	0.968	0.971	0.992
Floreana	0.041	0.026	0.032	—	0.979	0.956
San Cristóbal	0.012	0.018	0.029	0.021	—	0.970
Eden	0.032	0.063	0.008	0.045	0.031	—

*Presented for each island-pair are Nei's [42] unbiased genetic identity (above diagonal) and distance (below diagonal) estimates.

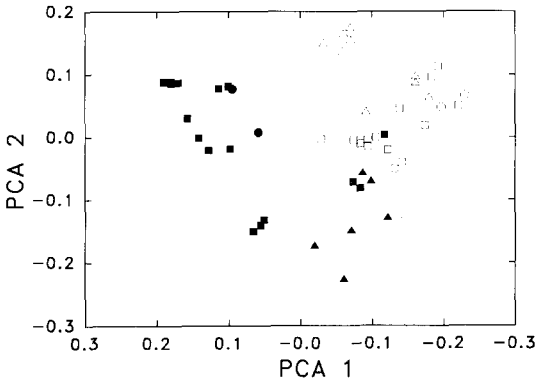

FIG. 2. PRINCIPAL COMPONENT ANALYSIS OF 51 *GOSSYPIUM DARWINII* ACCESSIONS BASED ON THE COVARIANCE MATRIX OF GENE FREQUENCIES. The first two axes explain 27.7 and 16.0% of the total variance, respectively. Closed squares=Isabela; open squares=San Cristóbal; closed circles=Eden; open circles=Gardner; closed triangles=Floreana; open triangles=Santa Cruz.

accessions from San Cristóbal, Santa Cruz and Gardner, and their relatively large distance from Isabela and Eden populations. Accessions from Floreana had intermediate values of *I* with all other islands (range=0.956–0.979). These data are closely paralleled using alternative methods of genetic similarity estimation (e.g. Rogers' distance, data not presented). Phenograms produced from average linkage cluster analysis, using either Rogers' or Nei's genetic distance, result in a topology that is fully consistent with island relationships suggested by both PCA and genetic identity computations (data not presented).

Discussion

Genetic variation in Gossypium darwinii

To understand the amount and distribution of genetic variation in *G. darwinii*, it is useful to provide an appropriate comparative context. Levels of allozyme variation in plant species and populations have been summarized in several reviews, the most recent by Hamrick and Godt [15], who compiled summary statistics of genetic variation in 449 plant species. Although only a small number of island endemics have been surveyed for electrophoretic variation, the available data indicate that island plants exhibit low levels of genetic variation (e.g. *Tetramolopium* [16], *Bidens* [17], *Dendroseris* [18], *Eichhornia paniculata* [19], *Lycopersicon cheesmanii* [20] and *G. klotzschianum* [4]). The paucity of genetic variation in island populations is generally attributed to stochastic factors, particularly founder events associated with island colonization (genetic bottlenecks) and genetic drift in small populations. Reductions in heterozygosity and allelic richness associated with bottlenecks [21] are expected to occur not only during the initial immigration to an island chain, but are also anticipated as populations migrate throughout an archipelago. Because numerous life-history features impact levels of allozyme variation [15], the most clear-cut examples of the negative effects of founder events and isolation on genetic variation are those where continental congeners or progenitors have been compared with island derivatives [4, 19].

Low levels of genetic variation need not be expected in all island plant populations, however, as pointed out by Witter and Carr [22] in their study of 18 species of Hawaiian silverswords (*Dubautia* spp. and *Wilkesia* spp.). Averaged across species, estimates of A (mean number of alleles per locus), P (percentage polymorphic loci), and H (mean panmictic heterozygosity per locus) were 1.29 (range=1.1–1.7), 24.3 (10.0–40.0), and 0.075 (0.013–0.124), respectively. These means are remarkably similar to those observed in the present study for *G. darwinii* (A=1.34, P=27.1, H=0.062; Table 3).

Although these values are only about half of the average for all plant species [15], they are considerably higher than the majority of other insular plant populations studied to date.

Witter and Carr [22] suggested that two factors promote greater variability in Hawaiian silverswords than in the other insular genera studied; (i) population size (which is often higher than in *Dendroseris*, *Bidens*, or *Tetramolopium*), and (ii) age since colonization (which may be more recent in the aforementioned genera). With respect to *G. darwinii*, population sizes may be huge; on some of the larger islands, populations are often continuous for many kilometers along the coastline and inland, although individual plants may be widely scattered [23]. These large population sizes would tend to retard the decay of genetic variability due to drift, thus promoting the retention of allelic diversity. Estimating the time since the ancestor of *G. darwinii* colonized the Galapagos is an uncertain undertaking because only indirect evidence is available. We would point out, however, that the species is widespread and morphologically variable throughout the archipelago, suggesting a relatively ancient dispersal to the islands. Most estimates of the age of the islands range from three to four million years [24, 25], setting the upper limit at something less than this. Also, tetraploid *Gossypium* has been estimated to have originated between one and two million years ago, based on nucleotide sequence divergence data derived from restriction site analysis of chloroplast genomes [26].

Inter-island divergence patterns

Gossypium darwinii exhibits considerable variability in a large number of vegetative and floral features [5, 23; unpublished observations]. Stephens and Rick [23] considered the range of variation to be greater in *G. darwinii* than for any other wild species in the genus (although they only had extensive field experience with New World species). It is not surprising, therefore, that Stephens and Rick noted distinct morphologies in populations from different islands, "as if certain combinations of genes had become randomly fixed in different small isolates."

This pattern is reflected in the allozyme data of Tables 2–5. Relatively few, mostly low-frequency, variants are restricted to single islands, but different suites of high-frequency alleles characterize each island. As a result, gene frequencies at the majority of loci are heterogeneous among islands (Table 4), and almost half of the total variation arises as a consequence of inter-island differentiation (G_{ST}=0.43; Table 4). Populations on different islands appear to be effectively reproductively isolated, suggesting a process of dispersal and colonization (accompanied by severe genetic bottlenecks) followed by population recovery and range expansion.

Perhaps the most surprising results of the allozyme analysis are the relationships revealed by PCA (Fig. 2). Projection of populations onto the first two principal components (accounting for 43.7% of the total variance) resulted in a general pattern of population clustering according to island of origin. Moreover, the two-dimensional depiction of relationships among islands approximates their geographic relationships, i.e. the first principal component approximates a west–east axis and the second component roughly corresponds to a north–south axis. Accordingly, PCA essentially generated a crude map of the Galapagos Islands (cf. Figs 1 and 2). The most notable exception involves the two populations from Eden, a small island near Santa Cruz. Rather than clustering near Santa Cruz populations, these accessions were grouped with the primary Isabela cluster, suggesting this latter island as the source of the founding propagules.

Few plant species have been demonstrated to exhibit such a striking relationship between geographic and genetic distance [e.g. 27], and to our knowledge no similar data exist for any island species. In fact, in the several species where this relationship has been explicitly tested [e.g. 19, 20], no significant associations were detected. The implication for *G. darwinii* is that colonization of the archipelago proceeded primarily via a "nearest-neighbor" or "stepping-stone" pattern of dispersal.

Relationships between G. darwinii *and* G. barbadense

A large number of allozyme studies have been conducted on progenitor-derivative species pairs, or species thought to have diverged relatively recently. In many genera [tabulated in 28], several predictions [29] regarding allelic distribu-

tion and levels of allozyme variation have generally been corroborated. These data, predicated on the assumption that speciation events will involve extreme genetic bottlenecks, have demonstrated that in most progenitor-derivative situations: (i) the allelic profile of the derivative usually represents a subset of that observed in the progenitor, with few if any unique alleles (almost always five or fewer); (ii) those unique alleles possessed by the derivative are usually rare or low-frequency variants; (iii) estimates of genetic diversity (A, P, H) are typically lower in the derivative; and (iv) interspecific genetic identities are equivalent to or only slightly lower than intraspecific estimates (typically 0.90 or above). A relevant example is provided by the Galapagos endemic *G. klotzschianum*, which apparently has been derived from the Mexican species *G. davidsonii* [4]: the alleles detected in *G. klotzschianum* represent a subset of those found in *G. davidsonii*, and *G. klotzschianum* is genetically depauperate compared to its progenitor; only three rare alleles were unique to *G. klotzschianum*; the interspecific Nei's genetic identity is 0.87, a value within the range of conspecific populations of many species [30].

It is informative to contrast these data for *G. davidsonii* and *G. klotzschianum*, which fit the progenitor-derivative model, with the allozyme data for *G. barbadense* and *G. darwinii*, which do so only with difficulty. As indicated by the frequency data of Table 2, *G. darwinii* and *G. barbadense* share common alleles at many loci; indeed, there is no locus that completely distinguishes the two species (although they are nearly fixed for alternative alleles at *Adh2*, *Arg2*, *Aco3*, and *Leu1*. Accordingly, the interspecific Nei's genetic identity is 0.949, a value typical of conspecific plant populations [although a far larger number of loci (59) were scored in the present study, including 36 monomorphic loci, potentially biasing this estimate upwards]. However, each species exhibits a large number of unique alleles: 12 were detected in *G. darwinii* (51 accessions sampled) vs 21 in *G. barbadense*, which had over double the sampling intensity [6]. This large amount of allelic novelty has not been reported in any derivative of a documented progenitor-derivative species-pair. Also, five of the 12 unique *G. darwinii* alleles (*Adh2-1*, *Enp2-5*, *Arg2-0.5*, *Leu1-3*, *Pgm6-9*) have moderate to high frequencies. Moreover, *G. darwinii* does not appear to be appreciably more genetically depauperate than *G. barbadense*, and panmictic heterozygosity per locus is nearly identical in both species (Table 3). These data suggest that, rather than viewing *G. barbadense* and *G. darwinii* as progenitor and derivative, it may be more appropriate to consider them as descendants from a common ancestor.

It is necessary to consider the relevance of the allozyme data to the concept of *G. darwinii* as a distinct species, inasmuch as its status has been uncertain ever since Darwin's specimens were subjected to taxonomic study. Although Watt [7] considered it a distinct species, both Robinson [31] and Stewart [32] included it within *G. barbadense*. As previously noted, several later authors [5, 10, 11] relegated *G. darwinii* Watt to varietal status within *G. barbadense* L. [as *G. barbadense* var. *darwinii* (Watt) J. B. Hutch.]. This taxonomic ambiguity emphasizes the close affinity between *G. darwinii* and *G. barbadense*, a view supported by a considerable body of additional genetic and molecular evidence, as well as by the allozyme data presented here. Harland [8], for example, noted their complete interfertility and the absence of any F_2 breakdown (a high frequency of non-germinable seed, moribund seedlings, or other aberrant recombinant types), a phenomenon often observed in the F_2 generation between *Gossypium* species [33]. The two taxa also share similar arrays of flavonoid compounds [34] and comprise a distinct, monophyletic clade based on phylogenetic analysis of restriction site mutations in the plastid genome [26].

These morphological and genetic similarities and the high interspecific genetic identity may be taken as evidence supporting the recognition of the Galapagos endemic as a variety of *G. barbadense*. There are, however, notable differences between the two taxa in morphological characteristics [7, 9]. Our own observations of numerous field and greenhouse-grown accessions indicate that "pure" *G. darwinii* (accessions that lack detectable introgressed *G. barbadense* and *G. hirsutum* genes) is characterized by and delimited from *G. barbadense* by a suite of morphological features: (i) nearly exclusively trilobed leaves; (ii) "leaky" glands (anthocyanin pigmentation surrounding the lysigenous

cavities known among cotton researchers as "gossypol glands"); (iii) smaller capsules, seeds (0.025–0.035 g seed^{-1} vs 0.090–0.180 g seed^{-1} in *G. barbadense*), and bracts subtending the flowers; (iv) red stem and pulvinus coloration; (v) a more upright habit with numerous slender, ascending branches; and (vi) sparse, non-spinnable, khaki or brown fiber (usually less than 1.3 cm in length). *Gossypium darwinii* also tends to require longer to reach reproductive maturity (more than one year), and is less productive at maturity than *G. barbadense.* The two taxa also differ in anthocyanin genes [35], and two restriction sites in their plastid genomes [26]. Taken together, these morphological and genetic differences, and the high level of allelic novelty displayed by *G. darwinii,* bolster the argument that *G. darwinii* deserves recognition at specific rank.

Regardless of the taxonomic circumscriptions adopted, *G. barbadense* and *G. darwinii* clearly are relatively recent descendants from a common ancestor. Our viewpoint is to recognise that, although only an intermediate level of genetic differentiation has arisen, there is a clear geographic and reproductive isolation between the two taxa, as well as morphological and genetic discontinuity; therefore we concur with the most recent monograph of the genus [9] in recognizing the Galapagos endemic as a distinct species.

Interspecific introgression

It has been suggested that a significant proportion of the morphological variability observed in *G. darwinii* is the result of gene flow from primitive forms of cultivated *G. barbadense* introduced from the South American mainland [5, 23]. The primary evidence used to support this assertion is the presence, in some *G. darwinii* populations, of cultivated characteristics that are usually strongly contrasted in wild and domesticated plants, e.g. lint length, lint color, and capsule size, *Gossypium barbadense* has never been cultivated on an agricultural scale in the Galapagos, but small populations of door-yard (commensal) forms were grown for casual use by settlers during the 19th and 20th centuries. Prior to permanent colonization, the islands were also frequented by transient visitors during the 16th to 18th centuries. Stephens and Rick [23] suggested that traces of introgression are evident in populations of *G. darwinii* near formerly or presently settled areas of Isabela, San Cristóbal, and Floreana. They proposed that either accidentally or intentionally introduced *G. barbadense* became established as a spontaneous component of the flora, providing the opportunity for genetic interchange with extant wild cottons (*G. darwinii*).

Two types of interspecific introgression were potentially detectable with our allozyme surveys, i.e. from *G. hirsutum* and from *G. barbadense.* Ongoing studies of over 700 accessions of *G. barbadense* [6] and *G. hirsutum* [Wendel, unpublished data] have identified 14 allozyme loci that discriminate the two species. At each of these loci *G. barbadense* and *G. hirsutum* are fixed or nearly fixed for alternate alleles; this large number of loci provides a sensitive set of species-specific markers for the detection of inter-specific introgression. Seven of the 58 *G. darwinii* accessions examined in this study (AS892, CB3087, CB3097, CB3098, WB1207, WB1249, PW56) proved to be "contaminated" with one or more *G. hirsutum* alleles (*Arg2-1, Enp1-4, Idh1-2, Pgd1-3, Tpi7-4*). This interpretation is based on the presence in these accessions of common *G. hirsutum* alleles [Wendel, unpublished data] that are unknown in "pure" *G. barbadense,* except *Enp1-4* [6] and otherwise are not detected in *G. darwinii.* An alternative interpretation is that the shared alleles represent phylogenetically ancestral characters retained in *G. darwinii* and *G. hirsutum* and lost in *G. barbadense.* This alternative is considered unlikely for reasons detailed below.

Because *G. hirsutum* apparently has no history of cultivation in the Galapagos Islands, the presence of introgressed *G. hirsutum* alleles in *G. darwinii* raises an important question about the time, place, and manner of introgression. An obvious possibility is that *G. hirsutum* introduction occurred at some time in the past, but that vestiges of this introduction no longer remain. A second alternative is that these alleles became introduced during nursery propagation of accessions subsequent to their original collection. This alternative is likely for several accessions (e.g. CB3086) that were omitted from analysis on the basis of morphological characteristics (see Experimental). However, we view nursery contamination as an unlikely explanation for the

majority of other, putatively introgressant accessions (listed above); these accessions fail to display morphological characteristics of *G. hirsutum* that would be predicted under a scenario of recent hybridization and introgression.

Because of this, and for the reasons presented below, we propose a third explanation for the detection of *G. hirsutum* alleles in *G. darwinii*, i.e. that *G. hirsutum* alleles in *G. darwinii* result from *G. barbadense* introgression rather than *G. hirsutum* introgression, and that the particular *G. barbadense* populations involved experienced introgression from *G. hirsutum* at some earlier time in their evolutionary history. This hypothesis may at first appear convoluted, but it is supported by several observations: (i) the morphology of most putative introgressant *G. darwinii* accessions is skewed towards *G. barbadense* rather than *G. hirsutum* [5; unpublished observations]; (ii) many primitive or obsolete *G. barbadense* cultivars and Pacific Basin accessions, which are likely candidates as donors, carry these same introgressed *G. hirsutum* alleles [6]; (iii) all of the introgressed *G. hirsutum* alleles detected in *G. darwinii* were also detected in introgressed *G. barbadense* [6]; (iv) only *G. barbadense* has a documented history of introduction into the Galapagos Islands [5, 23], providing the critical opportunity for introgression; and (v) *G. hirsutum* alleles in *G. darwinii* are present at only seven of the 13 loci at which introgression could potentially be detected; two of these seven loci, *Arg2* and *Idh1*, accounted for 55% of the introgressant *G. hirsutum* alleles. Given the number of introgressant accessions identified, it would seen unlikely that relatively recent intermingling with *G. hirsutum* would result in such a strong skewing of introgressant loci. Although clearly speculative, these observations provide us with no apparent alternative to the hypothesis of a mediated transfer of *G. hirsutum* germplasm through introduced *G. barbadense.*

The second type of introgression that was potentially detectable using allozymes, i.e. transfer of "pure" *G. barbadense* alleles into *G. darwinii*, turns out to be more difficult to detect, due to the absence of suitable genetic markers that confidently discriminate the two species (Table 2). In addition, alternative explanations of allelic overlap must be considered, especially the phylogenetic possibility that shared alleles represent primitive characters retained from the common ancestor of both species. The best candidates for introgressed *G. barbadense* alleles would be those with distributions like *Arg2-4*, *Adh2-4*, and *Leu1-4*, which are common in *G. barbadense* but are rare to infrequent in *G. darwinii*. *Arg2-4*, for example, was not detected in *G. darwinii* outside of four suspected introgressant accessions (CB3120, PW37, PW44, PW53). Approximately half of the accessions containing *Adh2-4* and/or *Leu1-4* have morphological phenotypes that are aberrant for *G. darwinii*, these exhibiting varying degrees of *G. barbadense* influence (unpublished data). These three alleles were detected from Isabela, Floreana, and San Cristóbal (islands with histories of human habitation), but were absent from Santa Cruz, Gardner, and Eden. Moreover, human activity on Isabela has been greatest in the southern half of the island, precisely the region where the putative introgressant *G. barbadense* alleles are found.

Concluding remarks

We have discussed several factors of potential significance in generating the relatively high levels of allozyme diversity (for an island endemic) observed in *G. darwinii*, i.e. large population sizes, a relatively ancient colonization of the archipelago, and interspecific introgression from *G. barbadense.* We suggest that it is primarily this latter process that has caused uncertainty regarding the taxonomic status of *G. darwinii.* One might also postulate that relatively high diversity in *G. darwinii* arose as a consequence of multiple, independent colonizations from its continental progenitor. A likely mode of dispersal from the mainland is oceanic drift via the Humboldt Current, which sweeps northward along the west coast of S. America and then turns west, directly towards the Galapagos Islands; this, in fact, is undoubtedly a significant factor in the floristic affinities between the Galapagos and western S. America [1–3]. An alternative is avian dispersal, although birds have not been observed to eat *G.darwinii* or *G. barbadense* seeds. Stephens [36] has noted several aspects of *G. darwinii*'s biology that argue in favor of oceanic transport: (i) it usually grows within 3 m of the high-tide mark; (ii) seeds are capable of floating for at least 10 weeks, without any loss of buoyancy; and (iii) seeds immersed in 3.5 M

NaCl for 10 weeks show no detectable loss of viability. Thus, *G. darwinii* seeds would be capable of traversing the approximately 1000 km of ocean between mainland S. America and the Galapagos Islands. Whether this voyage was made more than once remains an open question. However, the stepping-stone pattern of inter-island colonization suggested from Fig. 2 is more consistent with a single rather than multiple introduction.

Although it is unlikely that avian dispersal has played a significant role in the origin of *G. darwinii*, Galapagos finches (Geospizinae) are known to use linted cottons in nest building [23]. At least one of our accessions (WB1227 from San Cristóbal) was originally obtained from an abandoned finch nest. It seems plausible, therefore, that birds have influenced both intra-island dispersal patterns and the migration of *G. darwinii* throughout the archipelago.

Experimental

Plants. A geographically representative selection of 58 accessions was assembled for electrophoresis from the USDA-ARS working germplasm collection at Maricopa, AZ (Table 1, Fig. 1). In selecting accessions, preference was given to those from the "PW" and "AS" expeditions (Table 1) due to their detailed locality information and recency of collection. Most accessions consisted of original field-collected seed (several accessions underwent a single renewal cycle), thus minimizing bias due to nursery contamination or possible drift from original genotypes. Many of the older "CB" and "WB" accessions displayed obvious signs of "improved" characteristics from cultivated cottons, presumably resulting from contamination or introgression by *G. hirsutum* or *G. barbadense* during many cycles of field evaluation and increase; these accessions were omitted from analysis. To maximize the possibility of discerning patterns of variability and regional relationships, accessions were chosen from six different islands (Table 1, Fig. 1), including Eden (two accessions), Isabela (19), San Cristóbal (15), Floreana (9), Santa Cruz (9), and Gardner (4). Detailed locality information is available [37, 38].

Electrophoresis and isozyme nomenclature. A portion of the cotyledons from germinating seedlings was used for starch gel electrophoresis. Seventeen enzyme systems were resolved using five different electrophoretic buffer systems; aspartate aminotransferase (AAT), phosphoglucose isomerase (PGI), endopeptidase (ENP), catalase (CAT), triose-phosphate isomerase (TPI), 6-phosphogluconate dehydrogenase (PGD), aconitate hydratase (ACO), alcohol dehydrogenase (ADH), NADP-isocitrate dehydrogenase (IDH), NADH-dehydrogenase (="menadione reductase", NAD), malate dehydrogenase (MDH), phosphoglucomutase (PGM), glutamate synthetase (GS), glutamate dehydrogenase (GDH), formate dehydrogenase (FDH), and both leucyl-specific (LEU) and arginyl-specific (ARG) forms of aminopeptidase. Details of sample preparation and gel and buffer composition were exactly as described in Percy and Wendel [6]. Enzymes were visualized using staining methods detailed in Wendel and Weeden [39]. Preliminary surveys indicated little or no variation within accessions; consequently, few individuals (an average of four) were analysed per accession.

Genetic interpretation of isozyme and allozyme phenotypes was based on observed patterns of variation, typical patterns of subcellular localization and gene expression in other plants, and knowledge of the quaternary structure of the protein products [reviewed in ref. 40]. Support for these interpretations comes from formal genetic analyses involving numerous interspecific and intraspecific F_2 and BC progenies [Wendel, unpublished data]. Loci encoding the most anodally migrating isozyme for each enzyme system were assigned the numerical designation 1, with additional loci numbered sequentially in order of decreasing electrophoretic mobility. Similarly, allozymes at each locus were given numerical designations in order of decreasing electrophoretic mobility.

Data analysis. Standard measures of genetic variability were computed for all accessions and various groups of accessions, including the proportion of polymorphic loci (P), the mean number of alleles per locus (A), and mean panmictic heterozygosity (H). Multivariate relationships among accessions were revealed with PCA using a covariance matrix derived from allele frequencies [41]. Recognition of accession groups based on these results allowed the computation of "island" gene frequencies. These were used in cluster analysis [41] and in apportioning genetic variation among regions [21]. This latter technique partitions total variation (H_T) into within-island and among-island components (H_S and D_{ST}, respectively); $G_{ST}(=D_{ST}/HT)$ is a measure of the proportion of total variation accounted for by regional differentiation. Homogeneity of gene frequencies among islands was tested by contingency chi-square analysis [14]. Genetic distance and identity statistics (*D* and *I*) were computed following Nei [42] and Rogers [43]. Many of the above computations were expedited by the computer programs BIOSYS (D. Swofford, Illinois Natural History Survey) and NTSYS (Exeter Publishing Ltd, Setauket, New York).

Acknowledgements—We thank P. Fryxell, E. Turcotte and F. Wilson for comments on the manuscript and T. Couch for technical assistance. This research was supported by the National Science Foundation (BSR-8619631 to J.F.W.)

References

1. Porter, D. M. (1983) in *Patterns of Evolution in Galapagos Organisms* (Bowman, R. I., Berson, M. and Leviton, A. E., eds), p. 33. Amer. Ass. Adv. Sci. San Francisco.
2. Porter, D. M. (1984) *Biol. J. Linn. Soc.* **21**, 243.
3. Porter, D. M. (1984) in *Key Environments—Galapagos* (Perry, R., ed.), p. 85. Pergamon Press, Oxford.
4. Wendel, J. F. and Percival, A. E. (1990) *Pl. Syst. Evol.* **171**, 99.
5. Hutchinson, J. B., Silow, R. A. and Stephens, S. G. (1947) *The Evolution of Gossypium.* Oxford University Press, London.
6. Percy, R. G. and Wendel, J. F. (1990) *Theor. Appl. Genet.* **79**, 529.
7. Watt, G. (1907) *The Wild and Cultivated Cotton Plants of the World.* Longmans, Green, London.

8. Harland, S. C. (1939) *The Genetics of Cotton.* Jonathan Cape, London.
9. Fryxell, P. A. (1979) *The Natural History of the Cotton Tribe.* Texas A&M Univ. Press, College Station.
10. Kearney, T. H. (1952) *Leafl. West. Bot.* **6**, 165.
11. Valíček, P. (1978) *Cot. Fib. Trop.* **33**, 363.
12. Wendel, J. F., Olson, P. D. and Stewart, J. M. (1989) *Am. J. Bot.* **76**, 1795.
13. Wendel, J. F., Stewart, J. M. and Rettig, J. (1990) *Evolution* (in press).
14. Workman, P. L. and Niswander, J. D. (1970) *Am. J. Hum. Genet.* **22**, 24.
15. Hamrick, J. L. and Godt, M. J. W. (1989) in *Plant Population Genetics, Breeding and Genetic Resources* (Brown, A. H. D., Clegg, M. T., Kahler, A. L. and Weir, B. S., eds), p. 43. Sinauer, Sunderland, MA.
16. Lowrey, T. K. and Crawford, D. J. (1985) *Syst. Bot.* **10**, 64.
17. Helenurm, K. and Ganders, F. R. (1985) *Evolution* **39**, 753.
18. Crawford, D. J. and Stuessy, T. F. (1987) *Syst. Bot.* **12**, 435.
19. Glover, D. E. and Barrett, S. C. H. (1987) *Heredity* **59**, 7.
20. Rick, C. M. and Fobes, J. F. (1975) *Evolution* **29**, 443.
21. Nei, M. (1987) *Molecular Evolutionary Genetics.* Columbia University Press, New York.
22. Witter, M. S. and Carr, G. D. (1988) *Evolution* **42**, 1278.
23. Stephens, S. G. and Rick, C. M. (1966) in *The Galapagos* (Bowman, R. I., ed.), p. 201. University California Press, Berkeley.
24. Cox, A. (1983) in *Patterns of Evolution in Galapagos Organisms* (Bowman, R. I., Berson, M. and Leviton, A. E., eds), p. 11. Amer. Ass. Adv. Sci., San Francisco.
25. Simkin, T. (1984) *Biol. J. Linn. Soc.* **21**, 61.
26. Wendel, J. F. (1989) *Proc. Natn. Acad. Sci. U.S.A.* **86**, 4132.
27. Wendel, J. F. and Parks, C. R. (1985) *Am. J. Bot.* **72**, 52.
28. Pleasants, J. P. and Wendel, J. F. (1989) *Am. J. Bot.* **76**, 1151.
29. Gottlieb, L. D. (1973) *Am. J. Bot.* **60**, 545.
30. Crawford, D. J. (1985) *Syst. Bot.* **10**, 405.
31. Robinson, B. L. (1902) *Proc. Am. Acad. Arts Sci.* **38**, 77.
32. Stewart, A. (1911) *Proc. Calif. Acad. Sci. Ser. 4* **1**, 7.
33. Stephens, S. G. (1950) *Bot. Rev.* **16**, 115.
34. Parks, C. R., Ezell, W. L., Williams, D. E. and Dreyer, D. L. (1975) *Bull. Torrey Bot. Club* **102**, 350.
35. Stephens, S. G. (1974) *J. Genet.* **61**, 128.
36. Stephens, S. G. (1958) *Am. Nat.* **92**, 83.
37. Percival, A. E. (1987) *The National Collection of Gossypium Germplasm.* Southern Coop Series Bull No. 321. Dept of Agr. Commun., Texas A&M University College Station.
38. Schwendiman, J., Ano, G. and Percival, A. E. (1983) *FAO/IBPGR Pl. Genet. Resources Newsl.* **64**, 33.
39. Wendel, J. F. and Weeden, N. F. (1989) in *Isozymes in Plant Biology* (Soltis, D. E. and Soltis, P. S., eds), p. 5. Dioscorides Press, Portland, OR.
40. Weeden, N. F. and Wendel, J. F. (1989) in *Isozymes in Plant Biology* (Soltis, D. E. and Soltis, P. S., eds), p. 46. Dioscorides Press, Portland, OR.
41. Sneath, P. H. A. and Sokal, R. R. (1973) *Numerical Taxonomy.* Freeman, San Francisco.
42. Nei, M. (1978) *Genetics* **89**, 583.
43. Rogers, J. S. (1972) *Studies in Genetics VII*, p. 145. Univ. Texas Publ. 7213. Austin, TX.